应激条件下的双重学习系统*

程一智[1]　吴寅[1]　李红[1,2,3]

（1 深圳大学心理与社会学院；
2 深圳市情绪与社会认知科学重点实验室；
3 深圳市神经科学研究院，深圳 518060）

摘　要　心理学、神经科学和行为经济学中许多研究一致认为，人的行为由双重学习系统控制，分别是反射性系统和反思性系统。前者是自动化的习惯系统，反应快速，不需要消耗认知资源，后者是反应较为缓慢的认知系统，需要消耗更多认知资源，但也更加灵活，可以有效应对外界环境的变化。这两种学习系统并行存在而又相互竞争，共同影响人的心理和行为。究竟是哪种学习系统对人的特定行为起着主导作用，又有哪些因素导致该系统对行为的支配，是近年来普遍受到关注的问题。过去研究者们采用导航学习、概率分类学习或工具性学习及其计算模型，从行为和大脑层面探讨双重学习系统在急性和慢性应激下的变化。通过回顾和分析这些研究，我们总结出应激导致个体向习惯行为转变的生理机制，即去甲肾上腺素和糖皮质激素在杏仁核的参与下与受体结合协同作用于双系统相关的脑区，并从该角度出发重新梳理和解释毒品成瘾的形成。未来需要重点关注个体基因差异与应激影响学习的关系，并采用多种研究手段以更好地揭示其背后的神经和内分泌机制。

关键词　应激；双重学习系统；反射性系统；反思性系统；毒品成瘾

许多研究都支持这样一种观点，人的行为过程可以置于一个谱系中，从快速的、基于习惯的、自动化的直觉加工，到缓慢的、目标导向的、受意识控制的理性加工，由此形成一个连续体(Daw, Niv, & Dayan, 2005)。该连续体的两端被普遍定义为反射性系统(reflexive system)和反思性系统(reflective system)，正常情况下它们在个体的学习和记忆过程中处于一种动态平衡。然而，大量研究表明，应激可以打破这两种学习系统的平衡，导致个体趋向于依赖反射性系统来指导行为，即产生习惯化倾向，并且有关动物的实验还观察到应激下相关脑区活动的改变(Yin & Knowlton, 2006)。一般认为，应激状态下个体的这种习惯反应是一种适应性机制，如在某些威胁场景中它有助于节省认知资源，避免犹豫不决或延迟行动导致的不良后果(Vogel, Fernandez, Joels, & Schwabe, 2016)。然而，这种行为却是以牺牲学习的灵活性为代价，使个体难以适应日益变化的外界环境(Dandolo & Schwabe, 2016)。另外，应激本身也可导致不良的生活习惯产生，譬如

收稿日期：2018-06-22

* 基金资助：国家自然科学基金项目(31671150)；广东省普通高校创新团队建设项目(2015KCXTD009)；广东省珠江人才计划项目(2016ZT06S220)；深圳市基础学科布局项目(2910424978)；深圳市基础学科布局项目(2910424978)；深圳市孔雀计划项目资助(KQTD2015033)；深圳市孔雀计划项目资助(016104926)；国家自然科学基金项目(31600923)

通讯作者：李红　E-mail: lihongszu@szu.edu.cn

吸烟、酗酒或不健康饮食，并最终可能增加有关精神疾病的患病风险(Juster, McEwen, & Lupien, 2010; McEwen, 2008; Schneiderman, Ironson, & Siegel, 2005)。研究应激对双重学习系统的影响以及个体差异，有助于理解应激作用于学习的认知和神经机制，提高个体处理应激的能力，同时也为相关神经或精神疾病的预防、诊断和治疗提供有益的参考。

1 双重学习系统

1.1 反射性系统与反思性系统

人类和其他有机体，经过多年漫长的进化，发展出在复杂的外界环境下趋利避害的本能。个体的这种行为是对刺激的简单反应，还是由受预期目标结果的驱使，一直是研究者们关注的问题。许多研究认为，个体受双重学习系统的控制，分别为反射性系统和反思性系统，它们共同存在于人类的大脑中，在不同的场景下起着相对独立或相互竞争的作用(Balleine & O'Doherty, 2010; Daw et al., 2005; Dickinson, 1985; Gilovich, Griffin, & Kahneman, 2002; Sloman, 1996)。

反射性系统，又叫习惯系统，建立在刺激与反应的联结。它的特点是自动化的，反应快速，不需要消耗太多认知资源，有利于节省精力应对其他复杂的环境，因此也是一种适应性的学习系统。然而，由于习惯系统不需要个体外显地思考即将采取的行动可能产生的结果，因此显得呆板而不灵活，无法适应外界多变的环境。反思性系统，又叫认知系统，建立在行为与结果的关联，需要个体外显地思考即将采取的行动所可能产生的结果。与反射性系统不同，它反应较为缓慢，需要消耗更多认知资源，但更加灵活，可以有效应对外界环境的变化。

有关反射性系统与反思性系统的研究大致有以下四种不同的实验传统：首先是导航学习(navigational learning)，跟托尔曼的早期工作密切相关。研究者提出了基于海马的空间学习(hippocampus-dependent spatial learning)和基于背侧纹状体的刺激反应学习(dorsal striatum-dependent stimulus-response learning)这对概念，研究个体行为在多大程度上受习惯学习系统的支配(Packard & Knowlton, 2002)。空间学习是指个体利用环境中的多重线索来构建“认知地图”(cognitive map)，刺激反应学习则指个体仅仅参照行为反应和单个刺激的联结来指导自己的行为。受该研究启发，概率类别学习(probabilistic classification learning)领域采用基于海马的认知系统(hippocampus-dependent cognitive system)和基于背侧纹状体的习惯系统(dorsal-striatum-dependent habit system)对双重学习系统做出区分。工具性学习(instrumental learning)提出的目标导向学习(goal-directed learning)，指导个体加工行为与结果的关联，其神经基础主要是眶额皮质(orbitofrontal cortex, OFC)，而习惯化学习(habitual learning)使得个体仅关注于反应与刺激的关系，忽略反应所导致的结果同行为之间的因果联结，主要涉及背内侧纹状体(dorsal medial striatum, DMS)。随着计算神经科学的兴起，研究者发展出工具性学习的计算模型(computational approach to instrumental learning)，提出基于模型的学习(model-

based learning)和无模型学习(model-free learning)，被认为是对目标导向学习和习惯学习的延伸。基于模型的学习依赖于前额叶(prefrontal cortex, PFC)，是一种高度灵活却又消耗认知资源的学习系统，它依据奖赏最大化原则帮助个体构建环境模型并指导后续的行为，而无模型学习涉及背侧纹状体，在它的控制下个体采用一种相对呆板不灵活的学习策略，只重复过去能得到奖赏的行为。

1.2 影响双重学习系统的因素

最早研究者把关注点集中在提供双重学习系统存在的证据，并从行为和大脑层面对它们进行分离，从而积累了大量的研究成果。近年来，人们又将研究兴趣转移到影响双重学习系统对行为控制的调节因素上来：哪种学习系统对人的特定行为起着主导作用，又有哪些因素导致该系统对行为的支配，越来越受到研究者的关注。

研究表明，为了使任务表现最优，通常情况下反射性系统与反思性系统需要同时参与到对行为的控制中(Shohamy, Myers, Hopkins, Sage, & Gluck, 2009)。然而在其他场景中，许多因素包括分心任务(Foerde, Knowlton, & Poldrack, 2006)、训练时间(Poldrack et al., 2001)以及应激事件(Schwabe & Wolf, 2012)均能影响和改变这两种学习系统对行为的主导作用。以训练时间为例，为了避免细菌感染，饭前洗手就是一种目标导向行为，它受反思性系统的控制。然而，当洗手的行为经过多次重复之后，反射性系统便开始对行为施加主导作用，如只要一进入厨房就洗手，个体不再思考洗手的意义。由此可见，正常的行为表现应该是双重学习系统处于动态平衡的结果，而过度训练导致了这两种系统的失衡。而应激也能影响双重学习系统，主要体现在灵活的认知策略减少，呆板的习惯反应增加(Schwabe et al., 2007; Schwabe & Wolf, 2009, 2012)。例如，经历急性应激后的被试虽然已经有饱腹感，却仍然对可获得某种食物的按键进行反应，表现出一种习惯化反应(Schwabe & Wolf, 2009)。应激可以导致个体皮质醇水平升高，从而促进反思性系统向反射性系统的转变(de Quervain, Roozendaal, & McGaugh, 1998; Schwabe, Schächinger, de Kloet, & Oitzl, 2010; Schwabe, Tegenthoff, Hoffken, & Wolf, 2013; Vogel et al., 2016)。本文将基于前人积累的成果，梳理应激影响双重学习系统的认知过程和神经机制。

2 关于应激

2.1 应激和应激反应

应激在日常生活中普遍存在，它对人们的心理健康有重大的影响。关于应激的定义，Koolhaas等人(2011)通过元分析的方法总结出，当外部刺激超出个体调节能力时应激就会产生。过去通常认为，引发应激的应激源具有不可控制性和不可预测性，但有新的观点指出，可以准确预测未来发生的事件也能诱发应激(Trapp, O'Doherty, & Schwabe, 2018)。

应激涉及两种生理系统，为交感神经-肾上腺髓质轴(Sympathetic-adrenal-medullary Axis, SAM axis)和下丘脑-垂体-肾上腺轴(Hypothalamic-pituitary-adrenal Axis, HPA axis) (Ulrich-Lai & Herman, 2009)。它们分别促进儿茶酚胺(包括肾上腺素和去甲肾上腺素)和糖皮质激素(glucocorticoids)的分泌，引发一系列外周神经反应，包括心率、血压和α淀粉酶分泌的变化。糖皮质激素可以进入大脑与糖皮质激素受体(glucocorticoid receptors, GRs)和盐皮质激素受体(mineralocorticoid receptors, MRs)结合，进而对认知功能相关的脑区如前额叶等产生影响。人类个体中的糖皮质激素为皮质醇(cortisol)，它可以提高血液中的血糖水平，为机体作出行为反应提供能量。

2.2 应激的诱发

根据应激的操作性定义，研究者设置应激源诱发被试的应激反应，预期皮质醇水平达到峰值时，让被试执行有关实验任务。Schwabe等人设计的带有社会性评价的冰水任务(Socially Evaluated Cold Pressor Test, SECPT) (Schwabe & Schachinger, 2018)，要求被试将手浸入冰水中，同时拍摄被试面部表情。生理和心理的双重威胁使得SECPT很大程度诱发应激反应，如皮质醇水平明显增加(Smeets et al., 2012)。简单的实验操作使它成为一种经济有效的应激诱发范式。特里尔社会应激测试也被广泛采用(Trier Social Stress Test, TSST) (Kirschbaum, Pirke, & Hellhammer, 1993)。被试需要进行公众演讲，然后完成数学运算，整个过程用摄像头记录，增加社会性评价因素。Ishizuka等人(2007)证明了该范式的有效性。而多人版本的TSST范式(The Trier Social Stress Test for Groups, TSST-G) (von Dawans, Kirschbaum, & Heinrichs, 2011)则提高了实验效率。TSST在应激诱发方面更为有效，但由于操作过程相对繁琐，研究者需要结合具体研究目的在两者之间做出权衡。

3 应激条件下的双重学习系统

过去研究者们在实验室场景中对实验对象诱发应激，并采用导航学习、概率类别学习、工具性学习及其计算模型，研究应激对双重学习系统的影响，同时试图揭示其效应背后的神经内分泌机制。从导航学习到工具性学习的计算模型，我们简单回顾了以下四个领域的研究，其中涉及问题起源、实验范式以及实验逻辑，然后分别梳理了各个领域中应激及其中介因素所导致的个体对某一学习系统的使用偏向。

3.1 导航学习

对导航学习的探讨可以追溯到20世纪初刺激反应理论和认知学习理论之争。刺激反应学习理论认为，白鼠在T型迷宫中学会以特定的方式对刺激进行反应，譬如在测试阶段学会右转然后获取食物，是由于食物强化了刺激与反应之间的联结。然而托尔曼却认为，白鼠学习的不是左转或右转的序列，而是在脑中形成有关迷宫分布的“认知地图”，在该地图的导航下找到食物。十字迷宫任务(Plus-Maze Task)结果支持了后者的观点。该迷宫具有南北两个相反的入口以及东西两种不同的目标地点。实验小鼠分别为空间学习组和反

应学习组，从不同入口进入迷宫寻找食物。空间学习组的食物目标地点不固定，而反应学习组食物总是固定在入口的右侧。实验发现，随着试次的增多，相比反应学习组，空间学习组的小鼠错误率很快下降，表现出空间学习能力(Tolman, Ritchie, & Kalish, 1946)。为了从行为层面区分空间学习和刺激反应学习，研究者以十字迷宫作为基础设计了导航学习任务。研究者将食物放置地点固定在十字迷宫的左侧，训练阶段让白鼠从北侧入口进入迷宫，保证所有白鼠均能成功学习通过右转从而得到食物。到了测试阶段，改变白鼠的入口方向，让其从南侧入口进。如果白鼠受到训练影响而重复相同的行动，如向右转，则体现出刺激反应学习。如果白鼠不受训练影响并能迅速做出相反的行动，如向左转，则体现出空间学习的控制。

第一个利用导航学习任务探讨应激导致的习惯化研究来自于Kim等人(2001)。他们对实验组小鼠施加电击来诱发小鼠的应激反应，然后让其执行莫里斯水迷宫任务(Morris water maze, MWM)。结果发现，应激组游向新平台的潜伏期更短，并且其中有一半小鼠采用了刺激反应策略而径直游向新平台，另一半小鼠游向新平台之前先通往原始平台地点，即采用了空间学习策略，但新旧平台之间的游泳距离显示其更多使用了刺激反应策略。为了验证以上动物研究结果，Schwabe等人(2007)设计出新的空间学习任务，考察急性应激对人类被试双重学习系统的影响。实验中向被试呈现一个三维房间模型，房间中央的方桌上有四张背面朝上的卡片，其中一张为胜利牌，位置固定且旁边有植物，四周墙面有门、窗和挂钟等线索，被试需要根据主试的反馈学习并指认四张卡片中的胜利牌。主试旋转房间并交换任意墙面，但胜利牌和植物位置固定且不告知被试。如果被试学习到胜利牌的位置总是跟植物相邻，则是一种刺激反应策略，如果是通过周围墙面的线索学习到胜利牌的位置，则是一种空间学习策略。研究发现，相比控制组，经历TSST15分钟后的应激组被试唾液皮质醇水平显著升高，并且在该任务中较多使用了刺激反应策略，较少使用了空间学习策略。这表明人类被试的研究结果大体上同Kim等人的动物研究结果一致，应激增强了基于背侧纹状体的习惯学习。研究者还指出，被试在应激下采用了刺激反应策略是一种适应性反应。近期一项脑成像研究重复了此结果(Vogel et al., 2017)。该研究采用虚拟导航任务以及与MRI兼容的SECPT，让被试右脚浸入0～2°C的冰水中3分钟，同时进行高难度心算任务，随后进行虚拟导航任务。行为结果发现，应激组被试主要依靠地标线索，反映了刺激反应学习，而控制组被试主要依赖地平线线索，运用了空间学习策略。脑成像数据表明，应激组被试的杏仁核激活更强，且与纹状体的功能连接更加活跃。另外，药物操作研究发现，摄入MR受体抑制剂的个体身上并未发现这些行为以及相关脑区的显著激活。该研究直接证明了MR受体和杏仁核在应激对学习系统的影响中发挥了重要的介导作用。除了急性应激，慢性应激以及大脑关键发展期发生的应激也同样导致个体趋向基于背侧纹状体的刺激反应学习(Schwabe, Bohbot, & Wolf, 2012; Schwabe, Dalm, Schachinger, & Oitzl, 2008)。譬如，如果胎儿期个体的母亲孕期经历了负性生活事件，个体长大后在虚拟导航任务中更多采用了呆板的刺激反应学习策略。这表明学习策略的选择也受到了生命早期应激的显著影响。

3.2 概率类别学习

受上述研究的启发，研究者提出了概率类别学习，个体需要通过多个试次的反馈学习建立线索和刺激之间的概率关系，从而对刺激进行分类。天气预报任务(weather prediction task，WPT)是被广泛采用的一种实验范式。在该任务中，有四张不同的卡片作为线索(cue)，有晴天和雨天两种天气作为结果(outcome)，线索卡片分别以特定概率预测天气，被试通过反馈习得每次呈现的一张或多张卡片的预测概率。同十字迷宫任务一样，完成该任务也涉及基于海马的认知系统和基于背侧纹状体的习惯系统，体现在个体在类别学习时使用的是外显策略还是内隐策略。使用外显策略的个体只依据某单一线索进行判断，如只要出现线索卡片 1 就判断是晴天，否则就判断为雨天。使用内隐策略的个体则先学习每张线索卡片的预测概率，然后综合同时出现的几张线索卡片形成线索模式，以此来对天气进行预测。研究者根据数学模型加以确认个体采取何种策略。有研究显示，采用前者时海马得到激活，采用后者时纹状体得到激活(Shohamy et al., 2004)。这表明外显策略主要有认知系统参与，而内隐策略则反映了习惯系统的控制。

同采用导航任务的研究一样，采用概率类别任务的研究也发现，应激减少了个体基于海马的学习，增加了基于背侧纹状体学习的偏向(Schwabe et al., 2013; Schwabe & Wolf, 2012)。Schwabe和Wolf(2012)运用功能核磁共振成像方法，采用天气预报任务，研究急性应激是否调节这两种学习系统在任务学习中的参与程度以及相关脑机制。实验组被试需要从线索卡片中学习并预测天气，控制组被试只需在相同的试次序列中按键回答屏幕上的线索卡片是否少于两张，因此不涉及学习。实验组被试完成SECPT，当应激结束后25分钟唾液皮质醇水平达到预期峰值时对所有被试进行核磁共振扫描，并让其执行类别学习任务。结果发现实验组被试更多采用多线索策略，但其任务成绩本身并没有受到应激影响。脑成像数据显示实验组被试纹状体激活，而控制组被试海马得到激活，这也首次在人类被试身上揭示出应激改变双重学习系统的脑机制。

3.3 工具性学习

由于空间学习可能涉及几何认知，因此空间导航任务在某种程度上无法推广到一般学习情境，于是有研究者把关注点放在工具性行为研究上。工具性学习可以由目标导向系统控制，其神经基础是眶额皮质与壳核(putamen)，反映的是个体对行动和结果的联结，并且行动的结果是个体所需要或渴望得到的。工具性学习也可以由习惯系统控制，它仅仅建立在刺激反应联结的基础之上，不受结果的影响，与之相关的主要脑区是尾状核(caudate) (Dickinson, 1985)。为了区分某一工具性行为是目标导向还是习惯化学习，许多研究发展出一套工具性学习范式，其中最经典的是结果贬值任务(outcome devaluation task) (Valentin, Dickinson, & O'Doherty, 2007)，它被普遍认为是证明习惯行为存在的金标准。训练阶段，被试学会在按键和食物奖赏

之间建立联结。每个试次中均有一按键能够高概率获得食物，另一按键低概率获得食物。贬值阶段，邀请被试吃下所得食物直到有饱腹感，消除获取该食物的动机。消退测试阶段，让被试重新进行按键反应。如果被试依然对高概率食物进行按键，则说明对结果贬值不敏感，其行为是受习惯系统的控制。如果被试不再对高概率食物进行按键，说明对结果贬值有敏感性，能根据动机的变化来指导自己的选择，受目标导向学习的控制。该任务最初是用在啮齿类动物脑损伤研究中，后来通过改进用于人类身上，研究者发现，眶额皮质和背内侧纹状体是目标导向行为的主要脑区，同时背侧纹状体是习惯行为的神经基础(Balleine & O'Doherty, 2010; Yin & Knowlton, 2006)。

有关小鼠应激的研究表明，相比非应激组，结果贬值之前受到应激的大鼠对结果价值变化不敏感，表现出更多的习惯化行为(Braun & Hauber, 2013)。为了研究应激对人类目标导向学习的影响，Schwabe和Wolf(2011)对人类被试使用了该工具性学习任务，他们发现，在测试阶段，相比控制组，应激组被试没有改变按键反应，表现出更多的习惯化行为。值得注意的是，该研究中应激的诱发是发生在贬值阶段之前，应激很可能对训练阶段的学习产生影响。为了避免这种混淆因素的干扰，研究者的另一项工作是将应激设置在贬值阶段之后，并重现了上述研究结果(Schwabe & Wolf, 2010)。在随后的研究中，给应激组的一半被试服用一种叫心安得的β肾上腺素阻断剂，以减少应激导致的生理反应，结果发现，用药组被试表现出目标导向行为(Schwabe, Hoffken, Tegenthoff, & Wolf, 2011)。结果价值改变之后的目标导向行为是一种灵活、具有适应性的行为，以上研究表明应激对该适应性行为具有损害作用，并且这种损害可能是以肾上腺激素为中介的。然而，这些结果难以解决应激究竟影响哪种学习系统的问题：是目标导向系统受到损害，还是习惯系统得到提升，又或是两者兼而有之。针对此问题，近期有研究发现，应激组被试在一些仅依赖习惯系统的试次中能够很好地完成任务，表明应激是通过对目标导向系统产生作用继而影响行为的灵活性，并没有导致习惯系统的损伤(Fournier, d'Arripe-Longueville, & Radel, 2017)。同基于海马或背侧纹状体的学习一样，工具性学习也受慢性应激和生命早期应激的影响，并且都是转向更加习惯化的反应(Dias-Ferreira et al., 2009; Seehagen, Schneider, Rudolph, Ernst, & Zmyj, 2015)。譬如，应激对认知功能的许多研究都是针对成年被试而展开，对个体生命早期的探讨还比较少。为此研究者设计了婴儿版的工具性学习任务，让婴儿学习按压按钮(Seehagen et al., 2015)。操纵实验使得学习阶段能导致亮灯和声响的两个按钮到了测试阶段不再有任何效应，即所谓的贬值，然后观察婴儿在该阶段的行为反应。对实验组婴儿采用三阶段应激，以便有效引起婴儿唾液皮质醇水平的升高(van Bakel & Riksen-Walraven, 2004)。结果发现，经历应激后的婴儿在测试阶段选择继续按压相同的按钮，尽管该按钮不再产生声音和光亮，而控制组婴儿按压习惯按钮后开始尝试另一个按钮，表明应激导致婴儿认知灵活性下降，面对变化了的环境时无法调整自己的行为。Soares等人(2012)使用脑成像研究发现，经历长时间医学考试准备的慢性应激者表现习惯化行为，同时伴随着尾状核、壳核结

构和功能的改变，并且，通过追踪研究发现，考试结束六个月后，原来的慢性应激组被试又重新表现出目标导向行为，说明慢性应激对大脑结构和功能的影响是可逆的，该结果对于应激的临床研究有一定的理论意义。以上研究表明，不论是急性应激、慢性应激还是生命早期应激，也不论是行为实验还是脑成像实验，均得出较为一致的结论，即应激加强了个体的习惯行为倾向。

3.4 工具性学习的计算模型

经过长期的发展，工具性学习领域积累了大量的研究成果，研究者在此基础上提出了习惯学习和目标导向行为的计算模型。Sutton和Barto(1998)受人工智能、最优控制理论的启发，提出强化学习理论(Reinforcement learning, RL)，并采用计算和建模方法研究个体如何通过一系列试误(trial and error)来学习所处的环境进而帮助个体在最大化获得奖赏的同时减少惩罚。最常用的范式叫系列决策任务(sequential decision task)，用来评估基于模型的学习(model-based)和无模型学习(model-free)两者之间的平衡。基于模型的学习者需要预测在每一特定状态下采取某种行动将产生的行动结果，采取该行动后比较既定的结果和先前的预测，得到预测错误(prediction error, PE)，根据PE来调整后续的预测和行动并更新有关该任务的内部模型：不同的行动如何导向不同的场景或结果，以及这些结果如何映射到价值。模型的建构有利于评估各种行动的潜在价值并指导个体做出后续选择。相反，无模型的学习者只关注行动的直接价值，维持过去有纯收益的行动，得到奖赏则增加后续该行动，得到惩罚则相应减少该行动。显然，基于模型的学习被认为是目标导向的，而无模型学习被认为是习惯化的(Daw, Gershman, Seymour, Dayan, & Dolan, 2011)。在系列决策任务中，每一试次的阶段一通常有两种不同的图形可供选择，每种图形均有固定的概率指向阶段二的状态。进入阶段二，被试同样需要在两种图形刺激中选择其一，然后得到金钱奖赏或无奖赏的反馈，图形对应金钱奖赏的概率在一定范围内浮动。习惯化的学习者总是重复选择阶段一能得到奖赏的图形，而不考虑阶段一到阶段二的转化是常见转化还是罕见转化。相反，目标导向的学习者则需要考虑该任务的转化结构，选择阶段一中能导致常见转化并最终得到奖赏的图形。研究者根据被试的选择，采用强化学习的计算模型得到被试基于模型学习和无模型学习各自的比重。

许多研究表明，在正常情境下，系列决策任务中学习者均同时表现出两种模型系统的控制(Daw et al., 2011; Gershman, Markman, & Otto, 2014; Glascher, Daw, Dayan, & O'Doherty, 2010)。基于模型的学习需要消耗认知资源，但具有灵活性，而无模型学习则节省资源，却牺牲了灵活性，这两种系统是并行存在而又相互竞争，共同控制个体的行为。但是，哪些因素可以影响这两种系统对行为的主导作用尚不清楚。研究者注意到，工作记忆容量的损耗可以减少个体对基于模型系统的依赖，但对无模型系统没有影响(Otto, Gershman, Markman, & Daw, 2013)，而先前的研究已经证明应激影响工作记忆以及前额叶中与工作记忆相关联的脑区(Lupien, Gillin, & Hauger, 1999; Qin, Hermans, van Marle, Luo, & Fernandez, 2009; Schoofs, Wolf, &

Smeets, 2009)，因此推测应激可以选择性地影响基于模型的学习同时不影响无模型的学习。于是Otto等人(2013)采用SECPT，在诱发被试应激10分钟后，让被试执行强化学习任务，同时采用工作记忆量表测查被试的认知能力。实验结果与假设一致，应激损害了个体对基于模型系统的利用，同时无模型系统不受影响。并且实验还发现，急性应激情境下，工作记忆能力较低的个体倾向于无模型学习，而工作记忆能力较高的个体可以有效避免急性应激导致的无模型学习偏向。而另外一项研究发现，应激反应的个体差异有效预测了基于模型学习的损伤情况，相比低慢性应激水平的被试，高慢性应激水平的被试在面对急性心理社会应激时表现出较低的基于模型的控制(Radenbach et al., 2015)。这些研究表明，工作记忆和慢性应激在急性应激导致的无模型学习倾向中起着调节作用，因此未来研究需要充分考虑HPA轴反应的个体差异问题，并在临床应用中有针对性地训练个体的认知能力和压力疏导，提高个体应对应激事件的能力。

需要指出的是，以上这四种研究传统，都是在学习系统二分法的的框架下发展而来，即它们都共同指向反射性系统和反思性系统。综上所述，我们认为，不论采用哪种学习类型和实验范式，也不论针对哪种被试类型，过去研究均得出了较为一致的结论，即应激导致基于大脑皮层的反思性系统向基于纹状体的反射性系统转变，并且这种效应是以神经递质和荷尔蒙(如去甲肾上腺素和糖皮质激素)为中介的。

4　应激影响双重学习系统的心理机制

应激影响双重学习系统的心理机制可以从巴甫洛夫—工具性转化(Pavlovian-Instrumental Transfer, PIT)的角度加以解释。巴甫洛夫条件反射也叫经典条件反射，指的是中性刺激与非条件刺激建立联结后可以成为条件刺激，从而诱发某一特定行为。例如，铃声与食物配对呈现，动物习得的标志是只要每次听到铃声就会流口水或朝向食物放置点。工具性条件反射与巴甫洛夫条件反射的主要区别在于前者需要个体做出主动的行为反应然后加以强化。例如，老鼠需要通过按压杠杆以获得食物，而按压杠杆的训练可以导致本文所关注的两种工具性行为，即习惯学习和目标导向行为。在PIT范式中，研究者对动物进行训练，首先是巴甫洛夫学习，如呈现光刺激后给动物提供食物，然后是工具性学习，即让动物学习通过按压杠杆来获得同样的食物奖赏。训练结束后研究者撤掉食物，把动物置于有杠杆的测试区，并为其呈现光刺激，结果发现动物按压杠杆的频率大大增加。该范式后来也被用于人类被试身上，研究条件刺激对工具性学习的影响及其神经机制。

尽管巴甫洛夫和工具性学习存在着很大的不同，然而许多研究表明，在许多条件下巴甫洛夫条件刺激却可以直接影响工具性条件学习(Rescorla & Solomon, 1967)。譬如，呈现能引起食欲的条件刺激，将促进按压杠杆以获得食物(Balleine, 2001; Hall, Parkinson, Connor, Dickinson, & Everitt, 2001)。类似地，呈现酒精相关的线索刺激，酒精成瘾者表现出对酒精奖赏的工具性行为(Rescorla & Solomon, 1967)。因此，巴甫洛夫条件刺激既然可以促进工具性学习，那么也就可以导致目标导向学习到习惯学习的转变。有研究为此推论提

供了证据，譬如，在动物模型中，研究者发现作为非条件刺激的应激源(如脚部电击或猎食者气味)导致了行为的习惯化转变(Packard, 2009)。

因此，我们可以这么理解，应激首先作为条件或非条件刺激，对个体的巴甫洛夫学习产生影响，通过PIT进而促进工具性学习，导致个体从目标导向到习惯学习的转变。

5 应激影响双重学习系统的神经机制

过去研究者通过fMRI和药物操纵法(包括受体激动剂和阻断剂)来研究应激影响双重学习系统的神经机制，也得出较为清晰的结论。在应激导致的习惯化转变中，发挥重要作用的激素包括去甲肾上腺素和糖皮质激素，相关的受体有MR受体，涉及的关键脑区有杏仁核以及前文已经提到的海马和背侧纹状体。

5.1 海马，背侧纹状体和杏仁核

有研究采用fMRI手段，发现经历应激的个体在概率类别学习中趋向习惯化学习，同时其海马体的神经活动减少，而海马体正是认知控制的关键脑区(Schwabe et al., 2013; Schwabe & Wolf, 2012)。为了探讨应激如何影响习惯系统脑区的激活，最近一项结合了EEG(脑电)和fMRI的研究发现应激组被试的背侧纹状体激活显著增强(Wirz, Wacker, Felten, Reuter, & Schwabe, 2017)。还有研究从功能连接的角度揭示出杏仁核在应激导致的习惯化反应中扮演着重要角色。具体来说，这些研究发现，应激增强了杏仁核—背侧纹状体的功能连接，同时降低了杏仁核—海马体的功能连接(Schwabe et al., 2013; Vogel et al., 2015; Vogel et al., 2017; Wirz et al., 2017)。而在较早的研究中，研究者在大鼠身上采用十字迷宫任务，并对大鼠杏仁核注射育亨宾(yohimbine, 一种致焦虑药物)，发现该药物导致大鼠向基于尾状核(背侧纹状体的一部分)的习惯学习转变(Packard & Wingard, 2004)。

5.2 去甲肾上腺素

值得一提的是，上述Packard的药物操纵研究中，育亨宾药物的致焦虑特性实际上一部分归因于其对去甲肾上腺素能系统的作用，因此也在某种程度上说明了去甲肾上腺素在应激导致的习惯化学习中所起的作用。为了验证这一点，Wirz等人(2017)考察了有特定基因差异的个体在应激下的概率类别学习，发现在应激条件下ADRA2B基因缺失变异体携带者相比非携带者表现出较低的习惯化偏向。而ADRA2B基因是人体中编码α2b型肾上腺素受体(α2b-adrenoceptor)的基因，缺少该基因的个体无法编码α2b型肾上腺素受体，因此也阻断了肾上腺激素对杏仁核的作用，从而减少了个体的习惯化反应。该研究证明去甲肾上腺素的正常功能也是不可缺少的重要一环。

5.3 糖皮质激素以及MR受体

糖皮质激素，如人体中的皮质醇被认为是应激和学习之间一个主要的调节因素。为了考察糖皮质激素的作用，研究者采用了药物操纵和fMRI相结合的研究手段，给实验组被试口服MR受体阻断剂，揭示了应激

与MR受体在影响杏仁核—背侧纹状体功能连接方面的交互作用(Vogel et al., 2015)。Schwabe等人(2010)对大鼠的背外侧纹状体和背内侧纹状体分别注射糖皮质激素，然后让大鼠执行空间迷宫任务，结果发现，只有背外侧纹状体接受注射的大鼠同历经过度训练的大鼠一样，表现出刺激反应学习的倾向。该研究首先表明，大鼠的背外侧纹状体构成了习惯学习的神经基础，在人身上对应的结构是壳核。其次，该研究揭示了糖皮质激素的作用以及表明了利用药物阻断该激素可以避免应激导致大鼠的习惯化偏向。另外两项在人类被试身上的研究也重复验证了糖皮质激素或MR受体的作用(Schwabe et al., 2013; Vogel et al., 2017)。

5.4 去甲肾上腺素和糖皮质激素的交互作用

以上研究表明了去甲肾上腺素和糖皮质激素各自在应激导致的习惯化学习方面的作用，但是并未能明确或直接揭示二者之间在影响行为控制的交互关系。为了考察它们的交互作用，研究者采用了药物操纵研究，给被试或单独口服氢化可的松(hydrocortisone)，或单独口服育亨宾，或两者同时服用，另外设置非用药对照组，然后让被试执行结果贬值任务(Schwabe, Tegenthoff, Hoffken, & Wolf, 2010)。结果发现，只有同时服用两种药物的被试在测试阶段表现出对结果贬值的不敏感性，即习惯化反应，而单独服用某一种药物和空白对照组被试的目标导向行为并没有受损。随后，研究者还在保留该研究设计的基础上增加fMRI，收集被试的脑成像数据，重复了行为结果的同时还发现，去甲肾上腺素和糖皮质激素的这种交互效应还跟内侧前额叶(medial prefrontal cortex, mPFC)和眶额皮质区域的活动减少显著相关(Schwabe, Tegenthoff, Hoffken, & Wolf, 2012)。而这些脑区正是构成目标导向学习的神经基础(Balleine & O'Doherty, 2010)。因此，去甲肾上腺素和糖皮质激素的协同作用还导致了目标导向学习系统的损伤。

总之，去甲肾上腺素和糖皮质激素在杏仁核的参与条件下协同作用于双系统相关的脑区，包括海马、背侧纹状体、内侧前额叶和眶额皮质，从而对空间学习或目标导向学习产生损害，导致个体从反思性系统控制到反射性系统控制的转变。

6 应激条件下的双重学习系统研究对毒品成瘾的启示

毒品成瘾(drug addiction)是一种长期的复发性疾病，以强迫性药物寻求和社会功能受损为特点，给个体及其家人和社会带来严重的负担。许多因素包括基因、人格、社会环境和生活事件可以导致成瘾(Brewer, Catalano, Haggerty, Gainey, & Fleming, 1998)，同时研究表明，应激可以影响个体用药动机、奖赏系统和药物药效，在毒品成瘾的形成和复发过程中起着关键的作用(Piazza & Le Moal, 1998; Sinha, 2001, 2008)。本文所探讨的应激对双重学习系统的影响，有利于加深对毒品成瘾形成和复发的认识，为进一步的临床诊断和治疗提供一定的启示(Packard, 2009; Packard & Goodman, 2012)。

Schwabe等人(2011)提出了应激影响毒品成瘾的认知模型，揭示了从自发的用药行为到随后的强迫性用药，再到成瘾复发背后的认知机制。研究者认为，在初期接触毒品阶段，个体虽然主要受目标导向系统控

制，但由于应激导致执行功能受损，使个体无法有效地抑制摄入毒品的行为。在摄入毒品之后，应激，包括急性应激和慢性应激，使得目标导向系统和习惯系统对工具性行为的控制产生失衡，导致个体趋向自动化的反应，因此表现出更多的毒品寻求和摄入行为，毒品成瘾由此形成。在毒品戒断期间，应激通过去甲肾上腺素和糖皮质激素的协同作用导致个体行为受习惯系统的控制，同时由于刺激反应联结已经形成，与毒品相关的线索刺激能唤起个体寻求和摄入毒品的反应，因此应激也就增加了个体成瘾复发的风险。研究者的另外两项研究中发现，β阻断剂或糖皮质激素拮抗剂可能有助于防止毒品成瘾的复发(Schwabe, Dickinson, et al., 2011; Schwabe & Wolf, 2011)。总之，应激对毒品成瘾和复发的影响是通过皮质醇水平的上升和促进双重学习系统的失衡为中介的。

应激对毒品成瘾的影响也可以通过应激影响双学习系统的心理机制即巴甫洛夫—工具性转化来理解。巴甫洛夫条件刺激可以促进工具性学习，导致目标导向学习到习惯学习的转变。同时有研究显示，经历早期应激个体的杏仁核较大，恐惧条件作用也更强(Tottenham & Sheridan, 2010)。由此可以推论，巴甫洛夫恐惧线索刺激也具备这种效应，尤其对于杏仁核功能异常的个体，习惯化倾向将会更加明显。因此，厌恶条件刺激可以改变个体的工具性行为，当有生命早期应激史的个体因某种原因想起过去的厌恶事件时，将更容易表现出习惯化行为，于是也更加具备物质滥用和毒品成瘾的风险(Dube, Felitti, Dong, Chapman, et al., 2003; Dube, Felitti, Dong, Giles, & Anda, 2003; Macleod et al., 2013)。

由以上两种模型可知，应激通过影响个体激素水平或大脑有关结构，进而导致双重学习系统的失衡，在毒品成瘾的形成和复发中起着关键的作用。因此，在毒品成瘾行为的临床治疗上，可以考虑在应激环节上进行对个体进行干预，如服用相关药物阻断皮质醇与MR受体的结合，或者通过认知训练提高个体应对应激的能力，从而间接有效地控制毒品成瘾的发展历程。

7 总结与展望

综上，我们梳理了急性应激或慢性应激条件下个体双重学习系统发生失衡的认知过程和生理机制，最后也从该视角出发重新解释毒品成瘾的形成和复发原因。在紧急或不可控情境下，急性应激反应可以短暂地提高个体的警觉和注意，促使个体趋向于运用习惯系统，加快行为反应，同时节省认知资源，这被认为是一种适应性机制(Vogel et al., 2016)。然而，当应激事件结束，这种应激导致的习惯化反应如果不能发生逆转，将严重影响个体学习的灵活性以及记忆的准确性(Dandolo & Schwabe, 2016)。甚至，长期的应激体验和不良习惯化行为还可以导致应激相关的心理和精神疾病，如应激创伤障碍和毒品成瘾(de Quervain, Schwabe, & Roozendaal, 2017; Goodman, Leong, & Packard, 2012)。事实上，应激状态下个体向习惯化学习的转变，短时间内也可以视为一种适应性行为，一旦学习系统失衡长期持续，将给个体带来负面影响。目前，应激条

件下双重学习系统的研究已经积累了不少成果，然而，我们认为未来的研究还需要围绕以下几点进行相应地改善。

首先是关于习惯系统的定义以及实验任务本身的可靠性。目前由于许多研究涉及不同的实验任务，对习惯学习普遍定义为一种非灵活的自动化行为，是建立在刺激和反应之间的联结，但是具体到每个任务中习惯学习的各项特征和属性并不十分明确，例如，习惯行为是先天的还是后天逐渐习得的，是否需要注意或意识的参与。即使采用同一种实验范式，许多脑成像研究之间发现激活的脑区也不尽相同。最近一项针对人类习惯学习的元分析发现，采用概率类别学习任务的研究有的报告了尾状核的激活，有的同时报告了尾状核与壳核的激活，并且体素峰值均位于相关结构靠前的部位。相反，采用迷宫导航学习的研究大多只发现了尾状核的激活，体素峰值部位比较靠后。值得一提的是，该研究还发现，采用结果贬值任务和强化学习范式的研究所得出的结论比较一致，普遍报告了侧壳核(lateral putamen)的激活(Patterson & Knowlton, 2018)。人类的壳核跟啮齿类动物中的背侧纹状体一样均为习惯行为的主要脑区，因此这些结果也表明，相对于导航学习和概率类别学习任务，结果贬值和强化学习任务更为有效地区分出习惯学习和目标导向学习两种系统，似乎是更为可靠的研究人类习惯行为的范式。未来研究可以结合前后这两类学习范式，有利于直接确认习惯反应的存在，同时也可以更加行之有效地区分出对贬值不敏感的个体或人群。

其次是有关个体差异对应激和双重学习系统之间关系的影响。应激从本质上理解，它是个体对外界环境的一种主观体验，这种主观性决定了群体的异质性，比如一种情景对某人来说是应激源，而对其他人来说却很可能是一种激励源。另外，如前文所述，个体工作记忆能力或者智商的差异也会对应激反应产生影响。有证据表明，认知差异有一部分是由基因决定的。例如，人体中编码α2b型肾上腺素受体的基因为ADRA2B基因，携带该基因的个体在应激情境下杏仁核激活更强(Cousijn et al., 2010)。而Wirz等人(2017)最近的研究采用概率类别学习任务，发现在应激条件下ADRA2B基因缺失变异体携带者相比非携带者表现出较低的习惯化偏向。未来可以借助这些研究的经验，关注个体差异尤其是基因水平的差异在应激和双重学习系统之间所起的调节作用，以期更好地理解应激影响双重学习系统背后的机制，并最终在临床应用中为应激敏感人群提供个性化治疗手段。

还有，需要考虑应激源的时间因素。时间因素包括应激任务潜伏期，应激源持续时间和个体早期应激三个因素(Lupien, McEwen, Gunnar, & Heim, 2009)。应激任务潜伏期，表示应激后、任务前的那一段时间。由于 GR 受体和 MR 受体对 HPA 轴影响下产生的糖皮质激素亲和性不同，MR 受体亲和性较大，与激素结合后迅速引发神经元兴奋，而 GR 受体亲和性较小，与激素结合后导致缓慢的对神经网络的抑制，因此，为了解释实验结果与 HPA 轴的相关性，对应激任务潜伏期的校正显得尤为重要。例如有研究表明，间隔几分钟的差异将足够影响应激对冒险行为的效应(Pabst, Brand, & Wolf, 2013; Salvador et al., 2017)。按应激源持续时

间，可将应激分为急性应激和慢性应激。有研究表明，在动物实验(Dias-Ferreira et al., 2009)和人类实验(McEwen, 2007)中，比起急性应激，慢性应激（即重复应激源的呈现）与决策相关脑区的结构变化存在相关。类似地，对于个体早期应激，随着个体生命周期内应激体验的不同，应激对决策的影响也不同(Lupien et al., 2009)。例如，经历早期生活压力的青少年容易表现出与决策有关的情绪或动机神经回路（例如，杏仁核、前额叶和纹状体）的改变(Tottenham & Galvan, 2016)。那么，应激对双重学习系统的影响是否也随着应激源时间因素的不同而有具体的变化，有待进一步的考察。

最后，未来需要采用多种手段进行研究。Dias-Ferreira和同事(2009)在小鼠身上考察了应激对目标导向行为的影响，他们发现，持续21天的应激(包括社会威胁、强制游泳和禁闭)之后，实验组小鼠在贬值阶段无法修正自己的习惯化按压杠杆的反应，并且实验还发现该组小鼠腹内侧前额叶(ventromedial prefrontal cortex, vmPFC)和背内侧纹状体均出现萎缩，同时背侧纹状体异常增大。该研究为应激导致行为改变的脑机制提供了直接证据。然而，由于科学研究的伦理问题，过去没有、将来也不会在人类被试身上展开类似人为致使脑损伤的实验。因此研究者采用了一种非侵入性的无创的研究方法，如经颅直流电刺激(transcranial direct current stimulation, tDCS)，研究发现，刺激背侧前额叶可以有效避免应激导致的工作记忆的损伤，揭示出背侧前额叶在应激和工作记忆之间的关系中发挥着关键作用(Bogdanov & Schwabe, 2016)。受该研究启发，后续研究可以采用类似的无创实验，采用脑功能成像技术，结合脑损伤病人的研究，进一步确定和区分应激条件下双重学习系统变化的大脑结构和功能。还可以展开更多的药物操纵研究，厘清应激的神经内分泌机制。譬如，通过前文所述，MR受体的作用已经很明晰，然而GR受体却很少得到探讨。已有的研究虽然首次表明了GR受体在慢性应激对工具性行为的作用(Gourley et al., 2012)，但是它在急性应激影响双系统平衡的作用方面却一无所知。因此，未来研究可以尝试填补这一空缺。

参考文献

Balleine, B. W. (2001). Incentive processes in instrumental conditioning. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 307–366). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology, 35*(1), 48–69.

Bogdanov, M., & Schwabe, L. (2016). Transcranial stimulation of the dorsolateral prefrontal cortex prevents stress-induced working memory deficits. *Journal of Neuroscience, 36*(4), 1429–1437.

Braun, S., & Hauber, W. (2013). Acute stressor effects on goal-directed action in rats. Neurobiology of Learning and *Memory, 20*(12), 700–709.

Brewer, D. D., Catalano, R. F., Haggerty, K., Gainey, R. R., & Fleming, C. B. (1998). A meta-analysis of predictors of continued drug use during and after treatment for opiate addiction. *Addiction, 93*(1), 73–92.

Cousijn, H., Rijpkema, M., Qin, S., van Marle, H. J., Franke, B., Hermans, E. J., . . . Fernandez, G. (2010). Acute stress modulates genotype effects on amygdala processing in humans. *Proceedings of the National Academy of Sciences of the United States of America, 107*(21), 9867–9872.

Dandolo, L. C., & Schwabe, L. (2016). Stress-induced cortisol hampers memory generalization. *Neurobiology of Learning and Memory, 23*(12), 679–683.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204–1215.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*(12), 1704–1711.

de Quervain, D., Schwabe, L., & Roozendaal, B. (2017). Stress, glucocorticoids and memory: implications for treating fear-related disorders. *Nature Reviews Neuroscience, 18*(1), 7–19.

de Quervain, D. J., Roozendaal, B., & McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature, 394*(6695), 787–790.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., . . . Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science, 325*(5940), 621–625.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 308*(1135), 67–78.

Dube, S. R., Felitti, V. J., Dong, M., Chapman, D. P., Giles, W. H., & Anda, R. F. (2003). Childhood abuse, neglect, and household dysfunction and the risk of illicit drug use: the adverse childhood experiences study. *Pediatrics, 111*(3), 564–572.

Dube, S. R., Felitti, V. J., Dong, M., Giles, W. H., & Anda, R. F. (2003). The impact of adverse childhood experiences on health problems: evidence from four birth cohorts dating back to 1900. *Preventive medicine, 37*(3), 268–277.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the United States of America, 103*(31), 11778–11783.

Fournier, M., d'Arripe-Longueville, F., & Radel, R. (2017). Effects of psychosocial stress on the goal-directed and habit memory systems during learning and later execution. *Psychoneuroendocrinology, 77*, 275–283.

Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). Retrospective revaluation in sequential decision making: a tale of two systems. *Journal of Experimental Psychology: General, 143*(1), 182–194.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*: Cambridge university press.

Glascher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron, 66*(4), 585–595.

Goodman, J., Leong, K. C., & Packard, M. G. (2012). Emotional modulation of multiple memory systems: implications for the neurobiology of post-traumatic stress disorder. *Reviews in the Neurosciences, 23*(5-6), 627–643.

Gourley, S. L., Swanson, A. M., Jacobs, A. M., Howell, J. L., Mo, M., Dileone, R. J., . . . Taylor, J. R. (2012). Action control is mediated by prefrontal BDNF and glucocorticoid receptor binding. *Proceedings of the National Academy of Sciences of the United States of America, 109*(50), 20714–20719.

Hall, J., Parkinson, J. A., Connor, T. M., Dickinson, A., & Everitt, B. J. (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *European Journal of Neuroscience, 13*(10), 1984–1992.

Ishizuka, K., Hillier, A., & Beversdorf, D. Q. (2007). Effect of the cold pressor test on memory and cognitive flexibility. *Neurocase, 13*(3), 154–157.

Juster, R.-P., McEwen, B. S., & Lupien, S. J. (2010). Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience and Biobehavioral Reviews, 35*(1), 2–16.

Kim, J. J., Lee, H. J. J., Han, J. S., & Packard, M. G. (2001). Amygdala is critical for stress-induced modulation of hippocampal long-term potentiation and learning. *Journal of Neuroscience, 21*(14), 5222–5228.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The Trier Social Stress Test - A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology, 28*(1-2), 76–81.

Koolhaas, J. M., Bartolomucci, A., Buwalda, B., de Boer, S. F., Fluegge, G., Korte, S. M., . . . Fuchs, E. (2011). Stress revisited: A critical evaluation of the stress concept. *Neuroscience and Biobehavioral Reviews, 35*(5), 1291–1301.

Lupien, S. J., Gillin, C. J., & Hauger, R. L. (1999). Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: A dose–response study in humans. *Behavioral Neuroscience, 113*(3), 420.

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience, 10*(6), 434–445.

Macleod, J., Hickman, M., Jones, H. E., Copeland, L., McKenzie, J., De Angelis, D., . . . Robertson, J. R. (2013). Early life influences on the risk of injecting drug use: case control study based on the Edinburgh Addiction Cohort. *Addiction, 108*(4), 743–750.

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiological Reviews, 87*(3), 873–904.

McEwen, B. S. (2008). Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators. *European Journal of Pharmacology, 583*(2-3), 174–185.

Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science, 24*(5), 751–761.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences of the United States of America, 110*(52), 20941–20946.

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision making: A few minutes make all the difference. *Behavioural Brain Research, 250*, 39–45.

Packard, M. G. (2009). Anxiety, cognition, and habit: a multiple memory systems perspective. *Brain Research, 1293*, 121–128.

Packard, M. G., & Goodman, J. (2012). Emotional arousal and multiple memory systems in the mammalian brain. *Frontiers in Behavioral Neuroscience, 6*, 14.

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the Basal Ganglia. *Annual Review of Neuroscience, 25*, 563–593.

Packard, M. G., & Wingard, J. C. (2004). Amygdala and "emotional" modulation of the relative use of multiple memory systems. *Neurobiology of Learning and Memory, 82*(3), 243–252.

Patterson, T. K., & Knowlton, B. J. (2018). Subregional specificity in human striatal habit learning: a meta-analytic review of the fMRI literature. *Current Opinion in Behavioral Sciences, 20*, 75–82.

Piazza, P. V., & Le Moal, M. (1998). The role of stress in drug self-administration. *Trends in Pharmacological Sciences, 19*(2), 67–74.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*(6863), 546–550.

Qin, S., Hermans, E. J., van Marle, H. J., Luo, J., & Fernandez, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological Psychiatry, 66*(1), 25–32.

Radenbach, C., Reiter, A. M., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H. J., . . . Schlagenhauf, F. (2015). The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology, 53*, 268–280.

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review, 74*(3), 151.

Salvador, A., Worbe, Y., Delorme, C., Coricelli, G., Gaillard, R., Robbins, T. W., . . . Palminteri, S. (2017). Specific effect of a dopamine partial agonist on counterfactual learning: evidence from Gilles de la Tourette syndrome. *Scientific Reports, 7*(1), 6292.

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology* (Vol. 1, pp. 607–628).

Schoofs, D., Wolf, O. T., & Smeets, T. (2009). Cold pressor stress impairs performance on working memory tasks requiring executive functions in healthy young men. *Behavioral Neuroscience, 123*(5), 1066–1075.

Schwabe, L., Bohbot, V. D., & Wolf, O. T. (2012). Prenatal stress changes learning strategies in adulthood. *Hippocampus, 22*(11), 2136–2143.

Schwabe, L., Dalm, S., Schachinger, H., & Oitzl, M. S. (2008). Chronic stress modulates the use of spatial and stimulus-response learning strategies in mice and man. *Neurobiol Learn Mem, 90*(3), 495–503.

Schwabe, L., Dickinson, A., & Wolf, O. T. (2011). Stress, habits, and drug addiction: a psychoneuroendocrinological perspective. *Experimental and clinical psychopharmacology, 19*(1), 53–63.

Schwabe, L., Hoffken, O., Tegenthoff, M., & Wolf, O. T. (2011). Preventing the stress-induced shift from goal-directed to habit action with a beta-adrenergic antagonist. *Journal of Neuroscience, 31*(47), 17317–17325.

Schwabe, L., Oitzl, M. S., Philippsen, C., Richter, S., Bohringer, A., Wippich, W., & Schachinger, H. (2007). Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Neurobiology of Learning and Memory, 14*(1), 109–116.

Schwabe, L., Schächinger, H., de Kloet, E. R., & Oitzl, M. S. (2010). Corticosteroids operate as a switch between memory systems. *Journal of Cognitive Neuroscience, 22*(7), 1362–1372.

Schwabe, L., & Schachinger, H. (2018). Ten years of research with the Socially Evaluated Cold Pressor Test: Data from the past and guidelines for the future. *Psychoneuroendocrinology, 92*, 155–161.

Schwabe, L., Tegenthoff, M., Hoffken, O., & Wolf, O. T. (2010). Concurrent glucocorticoid and noradrenergic activity shifts instrumental behavior from goal-directed to habitual control. *Journal of Neuroscience, 30*(24), 8190–8196.

Schwabe, L., Tegenthoff, M., Hoffken, O., & Wolf, O. T. (2012). Simultaneous glucocorticoid and noradrenergic activity disrupts the neural basis of goal-directed action in the human brain. *Journal of Neuroscience, 32*(30), 10146–10155.

Schwabe, L., Tegenthoff, M., Hoffken, O., & Wolf, O. T. (2013). Mineralocorticoid receptor blockade prevents stress-induced modulation of multiple memory systems in the human brain. *Biological Psychiatry, 74*(11), 801–808.

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience, 29*(22), 7191–7198.

Schwabe, L., & Wolf, O. T. (2010). Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. *Psychoneuroendocrinology, 35*(7), 977–986.

Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. *Behavioural Brain Research, 219*(2), 321–328.

Schwabe, L., & Wolf, O. T. (2012). Stress modulates the engagement of multiple memory systems in classification learning. *Journal of Neuroscience, 32*(32), 11042–11049.

Seehagen, S., Schneider, S., Rudolph, J., Ernst, S., & Zmyj, N. (2015). Stress impairs cognitive flexibility in infants. *Proceedings of the National Academy of Sciences of the United States of America, 112*(41), 12882–12886.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain, 127*, 851–859.

Shohamy, D., Myers, C. E., Hopkins, R. O., Sage, J., & Gluck, M. A. (2009). Distinct hippocampal and basal ganglia contributions to probabilistic learning and reversal. *Journal of Cognitive Neuroscience, 21*(9), 1821–1833.

Sinha, R. (2001). How does stress increase risk of drug abuse and relapse? Psychopharmacology, 158(4), 343–359.

Sinha, R. (2008). Chronic stress, drug use, and vulnerability to addiction. *Annals of the new York Academy of Sciences, 1141*(1), 105–130.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin, 119*(1), 3.

Smeets, T., Cornelisse, S., Quaedflieg, C. W., Meyer, T., Jelicic, M., & Merckelbach, H. (2012). Introducing the Maastricht Acute Stress Test (MAST): a quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses. *Psychoneuroendocrinology, 37*(12), 1998–2008.

Soares, J. M., Sampaio, A., Ferreira, L. M., Santos, N., Marques, F., Palha, J. A., . . . Sousa, N. (2012). Stress-induced changes in human decision-making are reversible. *Translational psychiatry, 2*(7), e131.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1): MIT press Cambridge.

Tolman, E. C., Ritchie, B. F., & Kalish, D. (1946). Studies in spatial learning. II. Place learning versus response learning. *Journal of experimental psychology, 36*(3), 221.

Tottenham, N., & Galvan, A. (2016). Stress and the adolescent brain: Amygdala-prefrontal cortex circuitry and ventral striatum as developmental targets. *Neuroscience and Biobehavioral Reviews, 70*, 217–227.

Tottenham, N., & Sheridan, M. A. (2010). A review of adversity, the amygdala and the hippocampus: a consideration of developmental timing. *Frontiers in Human Neuroscience, 3*, 68.

Trapp, S., O'Doherty, J. P., & Schwabe, L. (2018). Stressful Events as Teaching Signals for the Brain. *Trends in Cognitive Sciences, 22*(6), 475–478.

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience, 10*(6), 397–409.

Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience, 27*(15), 4019–4026.

van Bakel, H. J., & Riksen-Walraven, J. M. (2004). Stress reactivity in 15-month-old infants: links with infant temperament, cognitive competence, and attachment security. *Developmental Psychobiology, 44*(3), 157–167.

Vogel, S., Fernandez, G., Joels, M., & Schwabe, L. (2016). Cognitive Adaptation under Stress: A Case for the Mineralocorticoid Receptor. *Trends in Cognitive Sciences, 20*(3), 192–203.

Vogel, S., Klumpers, F., Krugers, H. J., Fang, Z., Oplaat, K. T., Oitzl, M. S., . . . Fernandez, G. (2015). Blocking the mineralocorticoid receptor in humans prevents the stress-induced enhancement of centromedial amygdala connectivity with the dorsal striatum. *Neuropsychopharmacology, 40*(4), 947–956.

Vogel, S., Klumpers, F., Schroder, T. N., Oplaat, K. T., Krugers, H. J., Oitzl, M. S., . . . Fernandez, G. (2017). Stress induces a shift towards striatum-dependent stimulus-response learning via the mineralocorticoid receptor. *Neuropsychopharmacology, 42*(6), 1262–1271.

von Dawans, B., Kirschbaum, C., & Heinrichs, M. (2011). The Trier Social Stress Test for Groups (TSST-G): A new research tool for controlled simultaneous social stress exposure in a group format. *Psychoneuroendocrinology, 36*(4), 514–522.

Wirz, L., Wacker, J., Felten, A., Reuter, M., & Schwabe, L. (2017). A deletion variant of the α2b-adrenoceptor modulates the stress-induced shift from "cognitive" to "habit" memory. *Journal of Neuroscience, 37*(8), 2149–2160.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience, 7*(6), 464–476.

Dual-learning systems under stress

CHENG Yizhi[1]; WU Yin[1]; LI Hong[1,2,3]

([1] *College of Psychology and Sociology, Shenzhen University*;

[2] *Shenzhen Key Laboratory of Affective and Social Cognitive Science*;

[3] *Shenzhen Institute of Neuroscience, Shenzhen 518060, China*)

Abstract: There is mounting evidence in psychology, neuroscience and behavioral economics to support the notion that human behavior is governed by dual-learning systems, namely, reflective, "cognitive" or reflexive, "habitual" system. The former one is performed automatically, responds quickly and does not consume cognitive resources. The latter one responds slowly and consumes more cognitive resources, but it is also more flexible and sensitive to the changes in the external environment. Both of these learning systems exist in parallel and compete with each other to jointly influence individual's mind and behavior. A widely concerned question in recent years is which system exerts dominant control over specific behavior and what factors determine whether reflective or reflexive system governs behavior. Over the past decades, researchers used navigation learning task, probabilistic classification learning or instrumental learning task and associated computational models to explore the changes of multiple learning systems under acute and chronic stress at both behavioral and neural levels. By reviewing these studies, we summarize the psychophysiological mechanism underlying the stress-induced bias toward habitual behavior, and reinterpret the causal relationship between this shift and drug addiction. Existing research shows that noradrenaline and glucocorticoids act through mineralocorticoid receptors and exert interactive impact on brain regions that subserve dual-learning systems, which is orchestrated by the amygdala. Future studies need to focus on the modulatory role of genetic differences in the effects of stress on learning, and use a variety of technical methods to elucidate its neuroendocrine basis.

Key words: stress; dual-learning systems; reflexive system; reflective system; drug addiction